

## CLAIMS

Sub A2  
1. A system for generating an operational assessment of a cache memory in a digital data processing system for respective cache memory sizes comprising:

- 3 A. an operational statistics gathering element for gathering operational statistics over a time  
4 interval, including a file information retrieval activity value and a extent of activity value for  
5 each file accessed during the time interval;
- 6 B. a cache miss prediction element for generating a cache miss prediction value in response to  
7 the operational statistics gathered by the operational statistics gathering element and a cache  
8 memory size value; and
- 9 C. a cache memory size adjustment element for adjusting the cache memory size in response to  
10 the cache memory size value generated by the cache miss prediction element for a selected  
11 one of said cache miss prediction values.

1 2. A system as defined in claim 1 in which the cache miss prediction element generates the cache miss  
2 prediction value based on a particular one of a plurality of cache memory management  
3 methodologies.

1 3. A system as defined in claim 2 in which one of said cache memory management methodologies is  
2 a FIFO (first-in/first-out) methodology, the cache miss prediction element generating the cache miss  
3 prediction value in accordance with:

-25-

$$1 = \sum_i \frac{A_i}{M + \frac{SA_i}{E_i}}$$

where "M" represents the cache miss prediction value, "S" represents the selected cache memory size value, "A<sub>i</sub>" represents the file retrieval activity value for a file "i," and "E<sub>i</sub>" represents the extent of activity value for the file "i."

4. A system as defined in claim 3 in which the cache miss prediction element determines the cache miss prediction value "M" using a binary search methodology over the interval

$$0 \leq M \leq \sum_i A_i = A$$

where "A" represents the total activity over the time interval.

5. A system as defined in claim 2 in which one of said cache memory management methodologies is a FIFO (first-in/first-out) methodology, the operational statistics gathering element further gathering a packet re-reference value indicating a number of times a portion of a file, identified as a packet, is referenced during the time interval, the cache miss prediction element generating the cache miss prediction value in accordance with:

-26-

$$1 = \sum_i \frac{\frac{A_i}{R_i} (R_i - 1 - H(i, S))}{M + \frac{S \left( \frac{A_i}{R_i} \right)}{E_i}}$$

where "M" represents the cache miss prediction value, "S" represents the selected cache memory size value, "A<sub>i</sub>" represents the file retrieval activity value for a file "i," "E<sub>i</sub>" represents the extent of activity value for the file "i," and "R<sub>i</sub>" represents the packet re-reference value for file "i."

6. A system as defined in claim 5 in which the cache miss prediction element determines the cache miss prediction value "M" using a binary search methodology over the interval

$$0 \leq M \leq \sum_i A_i = A$$

where "A" represents the total activity over the time interval.

7. A system as defined in claim 2 in which one of said cache memory management methodologies is an LRU (least-recently used) methodology, the operational statistics gathering element further

-27-

gathering a packet re-reference value indicating a number of times a portion of a file, identified as a packet, is referenced during the time interval, the cache miss prediction element generating the cache mis prediction value in accordance with:

$$1 = \sum_i \frac{\left( \frac{A_i T_i}{R_i} \right) M + \left( \frac{A_i S I}{R_i} \right)}{\left( S I + \frac{S A_i T_i}{E_i R_i} \right) M + \left( \frac{A_i I}{E_i R_i} \right) S^2}$$

where "M" represents the cache miss prediction value, "S" represents the selected cache memory size value, "A<sub>i</sub>" represents the file retrieval activity value for a file "i," "E<sub>i</sub>" represents the extent of activity value for the file "i," "R<sub>i</sub>" represents the packet re-reference value for file "i," and "I" represents the duration of the time interval.

8. A system as defined in claim 7 in which the cache miss prediction element determines the cache miss prediction value "M" using a binary search methodology over the interval

$$0 \leq M \leq \sum_i A_i = A$$

where "A" represents the total activity over the time interval.

-28-

9. A method for generating an operational assessment of a cache memory in a digital data processing system for respective cache memory sizes comprising the steps of:

- A. gathering operational statistics over a time interval, including a file information retrieval activity value and a extent of activity value for each file accessed during the time interval;
- B. generating a cache miss prediction value in response to the operational statistics gathered during the operational statistics gathering step, and a cache memory size value; and
- C. adjusting the cache memory size in response to the cache memory size value generated during the cache miss prediction step for a selected one of said cache miss prediction values.

10. A method as defined in claim 9 in which during the cache miss prediction step the cache miss prediction value based on a particular one of a plurality of cache memory management methodologies.

11. A method as defined in claim 10 in which one of said cache memory management methodologies is a FIFO (first-in/first-out) methodology, during the cache miss prediction step the cache miss prediction value being generated in accordance with:

$$1 = \sum_i \frac{A_i}{M + \frac{SA_i}{E_i}}$$

-29-

5 where "M" represents the cache miss prediction value, "S" represents the selected cache memory size  
 6 value, " $A_i$ " represents the file retrieval activity value for a file "i," and " $E_i$ " represents the extent of  
 7 activity value for the file "i."

1 12. A method as defined in claim 11 in which, during the cache miss prediction step, the cache miss  
 2 prediction value "M" being generated using a binary search methodology over the interval

$$0 \leq M \leq \sum_i A_i = A$$

3  
 T269X  
 4 where "A" represents the total activity over the time interval.

5 13. A method as defined in claim 10 in which one of said cache memory management methodologies  
 6 is a FIFO (first-in/first-out) methodology, the operational statistics gathering element further  
 gathering a packet re-reference value indicating a number of times a portion of a file, identified as a  
 packet, is referenced during the time interval, during the cache miss prediction step the cache miss  
 prediction value being generated in accordance with:

$$1 = \sum_i \frac{\frac{A_i}{R_i} (R_i - 1 - H(i, S))}{M + \frac{S \left( \frac{A_i}{R_i} \right)}{E_i}}$$

-30-

7 where "M" represents the cache miss prediction value, "S" represents the selected cache memory size  
 8 value, "A<sub>i</sub>" represents the file retrieval activity value for a file "i," "E<sub>i</sub>" represents the extent of activity  
 9 value for the file "i," and "R<sub>i</sub>" represents the packet re-reference value for file "i."

1 14. A method as defined in claim 13 in which, during cache miss prediction step, the cache miss  
 2 prediction value "M" being generated using a binary search methodology over the interval

$$0 \leq M \leq \sum_i A_i = A$$

4 where "A" represents the total activity over the time interval.

15. A method as defined in claim 10 in which one of said cache memory management methodologies  
 is an LRU (least-recently used) methodology, during the operational statistics gathering step a packet  
 re-reference value being further gathered indicating a number of times a portion of a file, identified  
 as a packet, is referenced during the time interval, the cache miss prediction element generating the  
 cache mis prediction value in accordance with:

$$1 = \sum_i \frac{\left( \frac{A_i T_i}{R_i} \right) M + \left( \frac{A_i S I}{R_i} \right)}{\left( S I + \frac{S A_i T_i}{E_i R_i} \right) M + \left( \frac{A_i I}{E_i R_i} \right) S^2}$$

27

-31-

7 where "M" represents the cache miss prediction value, "S" represents the selected cache memory size  
 8 value, "A<sub>i</sub>" represents the file retrieval activity value for a file "i," "E<sub>i</sub>" represents the extent of activity  
 9 value for the file "i," "R<sub>i</sub>" represents the packet re-reference value for file "i," and "I" represents the  
 10 duration of the time interval.

1 16. A method as defined in claim 15 in which during the cache miss prediction step the cache miss  
 2 prediction value "M" being generated using a binary search methodology over the interval

250X  
3

$$0 \leq M \leq \sum_i A_i = A$$

4 where "A" represents the total activity over the time interval.

Added  
A5